

University of Wollongong

Research Online

Faculty of Engineering and Information
Sciences - Papers: Part A

Faculty of Engineering and Information
Sciences

1-1-1993

Markov pyramid models in image analysis

Jennifer L. Davidson
Iowa State University

Noel A. Cressie
Iowa State University, ncressie@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/eispapers>



Part of the [Engineering Commons](#), and the [Science and Technology Studies Commons](#)

Recommended Citation

Davidson, Jennifer L. and Cressie, Noel A., "Markov pyramid models in image analysis" (1993). *Faculty of Engineering and Information Sciences - Papers: Part A*. 6027.
<https://ro.uow.edu.au/eispapers/6027>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Markov pyramid models in image analysis

Abstract

The use of statistical pattern recognition techniques in image processing has led to simplifying assumptions on the statistical interdependence of the pixel value of an image, which allow theoretical analysis and/or computational implementation to be achieved. For instance, the assumption of statistical independence of the values or that their joint distributions are multivariate normal, simplifies the analysis enormously. However, these results are very limiting in representing models for data, and do not allow for analysis of arbitrary spatial dependencies, in the data. One method for modeling two-dimensional data on a lattice array has been developed by Abend et al. called the Markov mesh model, and is a generalization of the familiar 1D Markov chain. The Markov mesh model allows the use of a class of spatial dependencies that is popular in many 2D data processing schemes, including image processing. One advantage of using this model is that it allows a computationally attractive implementation of statistical procedures involving joint and conditional probabilities. In this paper, we generalize Abend et al.'s results to a more comprehensive model, which we call the Markov pyramid model, using the concept of partial ordering. We present the necessary background for this model and show that Abend's model is a special case of our model. Finally, we present a simple application of our results to texture modeling.

Keywords

analysis, models, markov, image, pyramid

Disciplines

Engineering | Science and Technology Studies

Publication Details

Davidson, J. L. & Cressie, N. A C. (1993). Markov pyramid models in image analysis. Proceedings of SPIE - The International Society for Optical Engineering (pp. 179-190).

Markov Pyramid Models in Image Analysis

J.L. Davidson

Department of Electrical Engineering

Iowa State University, Ames, Iowa 50011

N.A.C. Cressie

Department of Statistics

Iowa State University, Ames, Iowa 50011

ABSTRACT

The use of statistical pattern recognition techniques in image processing has led to simplifying assumptions on the statistical interdependence of the pixel values of an image, which allow theoretical analysis and/or computational implementation to be achieved. For instance, the assumption of statistical independence of the values or that their joint distributions are multivariate normal, simplifies the analysis enormously. However, these results are very limiting in representing models for data, and do not allow for analysis of arbitrary spatial dependencies in the data. One method for modeling two-dimensional data on a lattice array has been developed by [Abend *et al.*, 1965], called the *Markov mesh* model, and is a generalization of the familiar one-dimensional Markov chain. The Markov mesh model allows the use of a class of spatial dependencies that is popular in many two-dimensional data processing schemes, including image processing. One advantage of using this model is that it allows a computationally attractive implementation of statistical procedures involving joint and conditional probabilities. In this paper, we generalize Abend *et al.*'s results to a more comprehensive model, which we call the *Markov pyramid* model, using the concept of partial ordering. We present the necessary background for this model and show that Abend's model is a special case of our model. Finally, we present a simple application of our results to texture modeling.

1. Introduction

A common representation of digital images involves an underlying rectangular $M \times N$ grid of pixel locations. When implementing an algorithm on a von Neumann or sequential machine, a scanning order of the grid must first be chosen. Two popular methods are the raster scan order of the array (from left to right and top to bottom), and the diagonal scan (along the diagonals from left to right and top to bottom). Each method induces an ordering of the grid points, where pixel location (i,j) is less than pixel location (h,k) if and only if location (i,j) is scanned before location (h,k) . Symbolically, we can represent this concept by $(i,j) < (h,k)$. In fact, the ordering so induced by each of these two scans is a *total* ordering of the set of pixel locations; that is, for every distinct pair of locations (i,j) and (h,k) , either $(i,j) < (h,k)$ or $(h,k) < (i,j)$.

One approach used in conducting statistical analysis of images is to assume statistical independence of the random variables (pixel values) representing the image. While it leads to simple distributional results, it is a very unrealistic assumption. Various other approaches that approximate the joint distribution from marginal distributions have been developed, although computation becomes infeasible except for trivial cases. The next natural probabilistic restriction to place on the pixel values is

Markovian dependence, where the pixel locations have a total order placed on them. For example, placing a total order on the pixel locations in the form of a space-filling curve (in a digital sense) results in spatial dependencies of the grid points in the array, so that the one-dimensional Markov property can be extended to two dimensions. The original concept of space-filling curves was first introduced by Hilbert in 1891 [Hilbert, 1891], and introduced by [Abend *et al.*, 1965] to provide a two-dimensional generalization of the one-dimensional Markov chain process. See Figure 1 for an example of such a curve. However, [Abend *et al.*, 1965] offered another model for data on a finite rectangular array that covered a larger class of spatial dependencies, called Markov mesh models. It turns out that this method of generalizing the Markov chain to two dimensions uses a particular partial ordering of the pixel locations. Informally, a partial ordering of a set is a relationship existing between (not necessarily all) pairs of elements in the set, satisfying reflexivity, antisymmetry, and transitivity. Not every pair of elements in the set needs to be related or ordered, hence the term “partial.” For example, the set of all subsets of a given set, under the relation of set inclusion, forms a partial order that is not a total order. Given minimal and reasonable assumptions concerning the conditional dependencies of the data, Abend showed that these conditional dependencies are only local. The computational advantages of these results make it feasible to fit conditional probability models to most images, regardless of size. In this paper we present a more generic model for such spatial dependencies using an arbitrary partial order. For reasons that will be apparent later, we call these models Markov pyramid models.

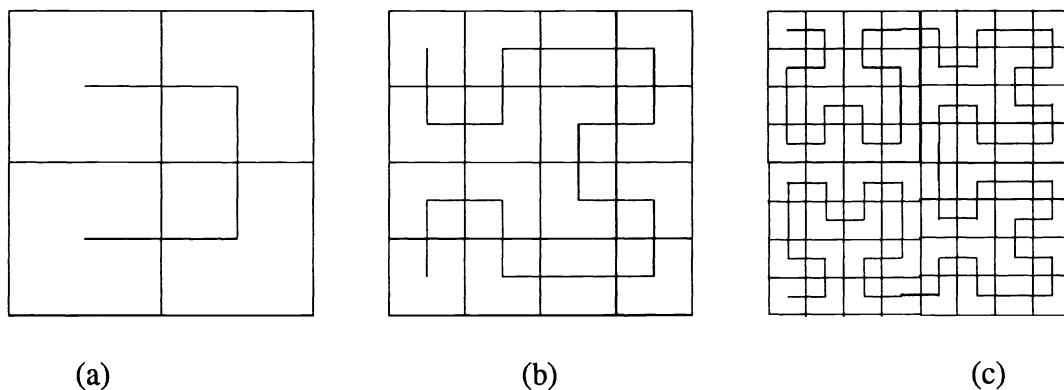


Figure 1. A space filling curve. (a) First iteration. (b) Second iteration. (c) Third iteration.

In the remainder of the paper, we present some statistical preliminaries, Abend *et al.*’s model and results, and then necessary concepts for defining the ordering relationship we use to generalize their model. We then define the Markov pyramid model, and show how Abend *et al.*’s Markov mesh model is a special case of ours. We conclude with an example of how our results might be applied to texture modeling. We remark that while our discussion is oriented towards image processing, our results apply to any finite set of variables whose sets possess a partial ordering.

2. Statistical Preliminaries

We shall assume some familiarity with basic probability and statistics, but for presentation purposes give a short introductory section on statistics and probability as we will be using it in this paper. Let $S \subset \mathbb{R}^n$ be a sample space, and let $A = \{U : U \subset S\}$ be a σ -field of Borel sets of S . Finally, suppose $P(\cdot)$ is a probability measure defined on A . Then the triple (S, A, P) is called a *probability space*. A *random variable* a is an A -measurable function

$$a : S \rightarrow \mathbb{R};$$

that is, the set $A_r = \{s \in S : a(s) \leq r\} \in A, \forall r \in \mathbb{R}$.

If $P(B) > 0$, then the *conditional probability of A given B* is

$$P(A | B) = \frac{P(A, B)}{P(B)}.$$

The *probability distribution function* of a random variable a with respect to the probability space (S, A, P) is the function F defined by

$$F(t) \equiv P(a \leq t) = P(\{s \in S : a(s) \leq t\}); \quad -\infty < t < \infty.$$

For the remainder of the paper, we will view our data as a finite collection of random variables, the domain of which is a subset of \mathbb{R}^2 . In section 3, the domain of the random variables is a finite subset of \mathbb{Z}^2 , while in section 5, the domain is an arbitrary finite subset of \mathbb{R}^2 .

3. The Markov Mesh Model

The Markov mesh model is defined for a finite set of random variables on an $M \times N$ grid. We make the following definitions:

1. Let $a_{i,j}$ be a random variable at location (i,j) in the $M \times N$ grid.
2. Let

$$X_{ij} = \{a_{h,k} : 1 \leq h \leq i \text{ and } 1 \leq k \leq j\}$$

denote the $i \times j$ array of random variables, and let

$$X_{MN} \equiv X = \{a_{i,j} : 1 \leq i \leq M, 1 \leq j \leq N\}$$

be the entire set of MN random variables on the $M \times N$ array.

3. Let

$$Z^{ij} = \{a_{h,k} : h < i \text{ or } k < j\}.$$

4. Let $X^{ij} = X \setminus \{a_{i,j}\}$ be the set X with the random variable $a_{i,j}$ deleted.

The sets described in items 2. and 3. are shown pictorially in Figure 2. The block with $a_{i,j}$ in it represents the pixel location with pixel value $a_{i,j}$.

The Markov mesh model assumes that the probability of $a_{i,j}$ conditional on $Z^{i,j}$ is equal to the probability of $a_{i,j}$, conditional on a limited set of neighbors immediately next to $a_{i,j}$ within the array X_{ij} . A third-order Markov mesh model satisfies

$$P(a_{i,j} | Z^{i,j}) = P(a_{i,j} | a_{i-1,j}, a_{i-1,j-1}, a_{i,j-1}),$$

for all i,j with $1 \leq i \leq M$ and $1 \leq j \leq N$. In [Abend *et al.*, 1965], the following result is obtained.

Theorem (Abend, Harley, and Kanal). The third-order Markov mesh condition gives a multiplicative decomposition for the marginal probability $P(X_{ij})$,

$$P(X_{ij}) = \prod_{h=1}^i \prod_{k=1}^j P(a_{h,k} | a_{h-1,k}, a_{h-1,k-1}, a_{h,k-1}).$$

In the above relation, if either subscript in the conditioning set is 0, it is assumed that the corresponding variable disappears from the conditioning set.

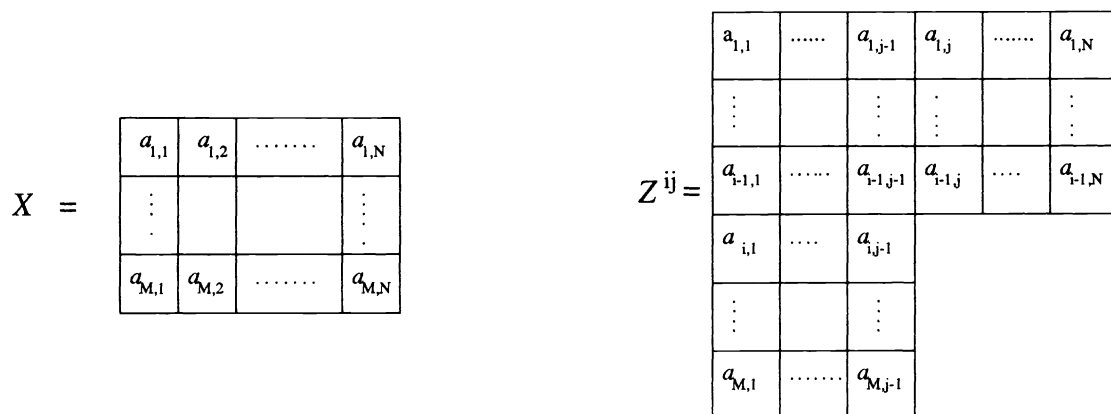


Figure 2. Pictorial descriptions of items 2. and 3. above.

In Figure 3 are some simple examples of local configurations that Abend *et al.*'s results can be extended to include. Figure 3(a) is the third-order Markov mesh model above; Figure 3(b) is a second-order Markov mesh model.



Figure 3. Two configurations giving rise to local conditional probabilities.

As a final remark, we note that Abend et al. use (although do not state) the following simple result.

Proposition. Let X be a set of random variables, and A, B and C be three subsets of X satisfying $C \subset B \subset A \subset X$. Suppose that $P(z | A) = P(z | C)$, for some $z \in X \setminus A$. Then $P(z | B) = P(z | C)$.

Proof: We show it for the discrete case; the proof for the continuous case is similar. First, note that

$$P(z | C) = P(z | B, A \setminus B) = \frac{P(z, A \setminus B | B)}{P(A \setminus B | B)}.$$

Let $A \setminus B = \{a_1, \dots, a_k\}$. Then we have

$$P(z | B) = \sum \sum \cdots \sum P(z, a_1, \dots, a_k | B)$$

where the sum is performed over all values that $a_i, i = 1, \dots, k$ can take; continuing,

$$\begin{aligned} P(z | B) &= \sum \sum \cdots \sum P(z, a_1, \dots, a_k | B) \\ &= \sum \sum \cdots \sum P(z | a_1, \dots, a_k, B) P(a_1, \dots, a_k | B) \\ &= \sum \sum \cdots \sum P(z | C) P(a_1, \dots, a_k | B). \end{aligned}$$

Because the summation is over all values for $\{a_i\}_{i=1}^k$ and $a_i \notin C$ for all i (since $\{a_1, \dots, a_k\} = A \setminus B$ and $C \subset B$), the term $P(z | C)$ can come outside the summation. Thus,

$$P(z | B) = P(z | C) \sum \sum \cdots \sum P(a_1, \dots, a_k | B) = P(z | C) \cdot 1,$$

as desired.

In their proof of the Theorem above, Abend et al. use this proposition for the case where X is the $M \times N$ array of random variables, $A = Z^{ij}$, $z = a_{i,j}$, $C = \{a_{i-1,j}, a_{i-1,j-1}, a_{i,j-1}\}$, and B is any subset of Z^{ij} containing C .

In section 4 we give the necessary background to the discussion of how to generalize this model.

4. Partial Orders

The main tool in our generalization is the use of a *partial order*. We shall show that the Theorem above is based on the use of a special partial order on the set of random variables X .

Let X be a set with a binary relation \leq . The relation \leq is a *partial order* for X if and only if it satisfies the following three properties:

1. (reflexivity) $a \leq a$, $\forall a \in X$;
2. (antisymmetry) $a \leq b$ and $b \leq a \Rightarrow a = b$, $\forall a, b \in X$;
3. (transitivity) $a \leq b$, $b \leq c \Rightarrow a \leq c$, $\forall a, b, c \in X$.

We call the pair (X, \leq) a *poset*. In a partial order, not every pair of elements need be related. For example, the subsets of a set under the relation of set inclusion is a partial order, because not every subset is related to every other subset by inclusion. An example of a set that has not only a partial order but a total order (where every element is related to every other element) is the set of real numbers under the relation of “less than or equal to.” A third example is given in Figure 4, which is a 7–element set whose partial orders are given by the arrows in the diagram. If $x \leq y$, then there is an arrow whose tail is on y and whose head is on x .

A fourth example, one of interest to us, is the partial ordering on the Cartesian product $\mathbb{Z} \times \mathbb{Z}$ defined by

$$(i, j) \leq (h, k) \Leftrightarrow i \leq h \text{ and } j \leq k.$$

In this example, the element $(h-1, k+1)$ is not related to (h, k) . We can apply this to the setting in Abend et al.’s Theorem. Let $X = \{a_{i,j} : 1 \leq i \leq M, 1 \leq j \leq N\}$, where $a_{i,j}$ is a random variable, and define a partial order on X by

$$a_{i,j} \leq a_{h,k} \Leftrightarrow i \leq h \text{ and } j \leq k.$$

(We have abused notation and have used the symbol \leq for the partial order as well as for the total order on the real numbers.) Here, the random variable $a_{h,k}$ is greater than or equal to every other element in $X_{h,k} = \{a_{i,j} : 1 \leq i \leq h, 1 \leq j \leq k\}$. Also notice that $a_{h-1, k+1}$ is not related to $a_{h,k}$. It is this partial order on $\mathbb{Z} \times \mathbb{Z}$ that is used by Abend et al. to prove their Theorem. Our main goals are to generalize this Theorem to an arbitrary partial order. That is, we seek a multiplicative decomposition for the joint distribution, assuming that the corresponding conditional probabilities $\{P(a_{ij} | Z^{ij})\}$ are spatially local.



Figure 4. (a) An arbitrary poset depicted as a directed graph (digraph). (b) The level sets $L^i, i = 0, 1, 2$, for the poset in (a).

By convention, we write $b \geq a$ to mean $a \leq b$, and $a < b$ means that $a \leq b$ but $a \neq b$. Similarly, $b > a$ means $a < b$.

We say an element a *covers* an element $b \neq a$ if $a > b$ and there does not exist a third element $z, z \neq a, z \neq b$, such that $a > z > b$. In other words, there is no other element z “in between” a and b . A partial order can be defined by stating the cover relationship between all appropriate elements in the poset. For Abend et al.’s example, by specifying that the element $a_{i,j}$ covers each of the elements $a_{i-1,j}, a_{i-1,j-1}, a_{i,j-1}$, we define the partial order as described above on $\mathbb{Z} \times \mathbb{Z}$. See Figure 5 for the

cover relationships between random variables. In Figure 5(a), we depict the partial order on elements in the array with the typical coordinate orientation. This does not give the notion of “up” and “down” as conveyed by the ordering, so in Figure 5(b) we have rotated the diagram to depict the “greater” elements on the top of the picture.

We define the set of all elements that a covers to be the set of *lower neighbors* of a , denoted by $LN a$. Thus,

$$LN a = \{b : a \text{ covers } b\}.$$

For example, for $a_{i,j}$, we have $LN a_{i,j} = \{a_{i-1,j}, a_{i-1,j-1}, a_{i,j-1}\}$. For a set $A \subset X$, we define the *cover* of A to be the set of elements of X not in A whose lower neighborhood lies entirely in A , and denote it $cov A$. Thus,

$$cov A = \{z \notin A : LN z \subset A\}.$$

Note that it is possible that $cov A = \phi$. In Figure 5, for example,

$$\begin{aligned} cov\{a_{2,2}\} &= \emptyset, \\ cov\{a_{1,1}\} &= \{a_{1,2}, a_{2,1}\}, \\ cov\{a_{1,1}, a_{1,2}, a_{2,1}\} &= \{a_{3,1}, a_{1,3}, a_{2,2}\}. \end{aligned}$$

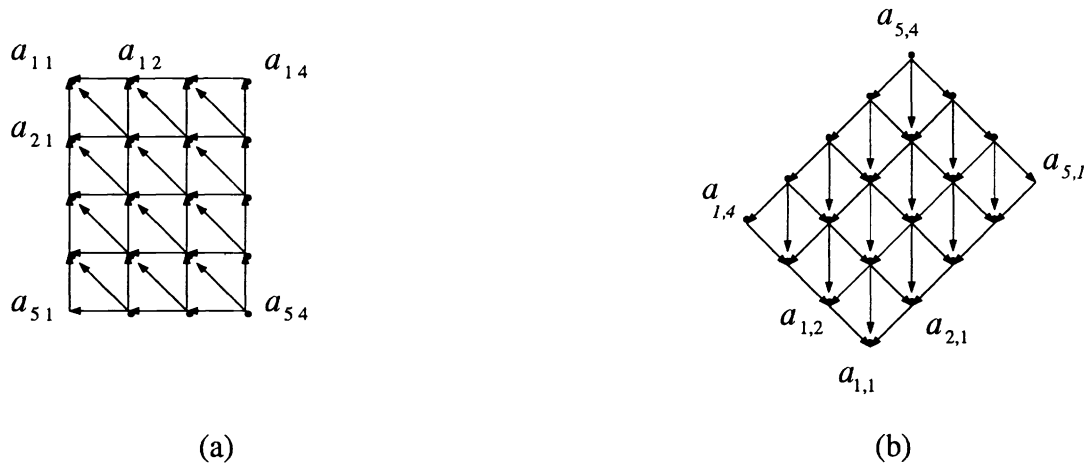


Figure 5. (a) A 5 x 4 array of random variables with the cover relation depicted by arrows. (b) Part (a) rotated so that the “biggest” element, $a_{5,4}$, is on “top.”

For an element a in X , we call the set of all elements that are strictly less than a , the *cone* of a . Thus,

$$\overline{cone} a = \{b : b < a, b \neq a\}.$$

If we depict the partial order by portraying the greatest elements on the top of the diagram with the arrows pointing downward, then all elements less than a particular element will fall beneath it. Figure 5(b) shows an example of such a cone, where

$$\text{cone } a_{5,4} = \{a_{i,j} : 1 \leq i \leq 4, 1 \leq j \leq 4\} \cup \{a_{5,1}, a_{5,2}, a_{5,3}\}.$$

We define the *closure* of the lower neighborhood of a , *closure* of the cover of A , and *closure* of the cone of a , as follows:

$$\begin{aligned}\overline{LN}a &= \{a\} \cup LN a \\ \overline{\text{cov}}A &= A \cup \text{cov}A \\ \overline{\text{cone}}a &= \{a\} \cup \text{cone } a.\end{aligned}$$

Again, for the example in Figure 5, we have

$$\begin{aligned}\overline{LN}a_{i,j} &= \{a_{i,j}\} \cup LN a = \{a_{i,j}, a_{i-1,j}, a_{i-1,j-1}, a_{i,j-1}\} \\ \overline{\text{cov}}\{a_{1,1}\} &= \{a_{1,1}\} \cup \text{cov}\{a_{1,1}\} = \{a_{1,1}, a_{2,1}, a_{1,2}\} \\ \overline{\text{cone}}a_{5,4} &= \{a_{5,4}\} \cup \text{cone } a_{5,4} = X_{5,4}\end{aligned}$$

Also, for any $a \in X$, the *dilation* of a with respect to the partial order is the set

$$\text{dil } a = \bigcup_{\{z: a \in \overline{LN}z\}} \overline{LN}z. \quad (1)$$

For our example, we have

$$\text{dil } a_{i,j} = \bigcup_{\{z: a_{i,j} \in \overline{LN}z\}} \overline{LN}z = \left\{ \begin{array}{ccc} a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} \\ a_{i,j-1} & a_{i,j} & a_{i,j+1} \\ a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} \end{array} \right\},$$

which is exactly the eight nearest neighbors of $a_{i,j}$ plus itself.

With every partial order on X there exists a unique directed graph, or digraph, where an edge between a and b exists if and only if a covers b or b covers a . If a covers b then there is a directed edge from a to b . For the proof, see for example [Harary *et al.*, 1965]. Because of this one-to-one mapping, we shall use the terminology for the partial order on X and for its corresponding digraph interchangeably. It is well-known that the digraphs that correspond to partial orders can have no sequence of distinct elements satisfying

$$a_1 \leq a_2 \leq \cdots \leq a_k = a_1,$$

that is, a *cycle* of elements. If this were to hold, by transitivity we would have $a_1 = a_2 = \cdots = a_k$, which contradicts the assumption of distinct elements. Thus, the digraphs that correspond to partial orders are *acyclic*.

We call an element x of X a *minimal element* if there is no other element z in X satisfying $x > z$. Similarly, a *maximal element* y is one for which there exists no other element z satisfying $y < z$. It can also be easily shown that the digraph associated with a finite poset X always has at least one minimal element and at least one maximal element. The set of minimal elements in X are denoted by X_{\min} , and the set of maximal elements by X_{\max} . In the example of Abend et al.'s partial order on the set X_{MN} ,

$$X_{\min} = \{a_{1,1}\}; \quad X_{\max} = \{a_{M,N}\}.$$

Every partial order on X defines a sequence of nonempty cover sets described in the following way [Harary *et al.*, 1965]. Let $L^0 = X_{\min}$, and recursively define $L^i = \text{cov}\left(\bigcup_{k=0}^{i-1} L^k\right)$. We call these sets the *level sets* for X . They are a special case of level sets as in [Harary *et al.*, 1965]. Our choice of level sets have the following properties:

Proposition 1. $L^i \cap L^j = \emptyset; i \neq j$.

Proposition 2. If $a \in L^i, i > 0$, then a covers some element in L^{i-1} .

Proposition 3. L^i contains only elements that are unrelated.

Proposition 4. If $a \in L^i, i > 0$, then $\text{cone } a \subset \bigcup_{k=0}^{i-1} L^k$.

In Abend et al.'s example, $L^0 = X_{\min} = \{a_{1,1}\}$, and $L^i = \text{cov}\left(\bigcup_{k=0}^{i-1} L^k\right) = \{a_{h,k} : h+k = i+2\}$. Thus, $L^1 = \{a_{1,2}, a_{2,1}\}$, $L^2 = \{a_{1,3}, a_{2,2}, a_{3,1}\}$, etc.

A *chain* is a sequence of distinct elements a_1, a_2, \dots, a_k such that every adjacent element in the sequence covers its neighbor; that is, a_i covers a_{i+1} , for $i=1, \dots, k-1$. The *length* of a chain with $k+1$ elements is k . The following result is true.

Proposition 5. Let X be a finite poset, and n be the length of the longest chain in X . Then

$$X = \bigcup_{k=0}^n L^k.$$

In Abend et al.'s example, $n = N+M-2$. Propositions 1 and 5 together state that these level sets partition X .

If we take an element $a \in X$, and look at its cone, we can partition its cone by intersecting the cone with each level set. This is an orderly way of categorizing the elements in cone a . Suppose that a is in level set L^i . Thus, if we define

$$L_a^k = L^k \cap \text{cone } a,$$

then

$$\text{cone } a = \bigcup_{k=0}^{i-1} L_a^k.$$

5. Markov Pyramid Models

We are now in a position to state our main result. Our Theorem 1 corresponds to Abend et al.'s Theorem. We assume that the poset X has no *singleton* points, that is, has no point that is not related to any other point. This is not an unreasonable assumption, since in many spatial problems we can assume that typically all points are statistically related, at least in some marginal way, to some other point. We call these models Markov pyramid models due to the pyramidal or cone structure in the digraph.

Theorem 1 (Davidson and Cressie). Let X be a finite poset with relation \leq and suppose that X has no singleton points. Let $\{L^k\}_{k=0}^n$ be the level sets as defined in Section 4. Let $a \in L^i \subset X$, and define $L_a^k = L^k \cap \text{cone } a$ as above, $k=0, \dots, i-1$. Suppose that

$$P(z \mid Y_z) = P(z \mid LNz), \forall z \in X,$$

where Y_z satisfies $LNz \subset Y_z \subset \text{cone } z \cup \{b : b \text{ and } z \text{ are not related}\}$. Then

$$P(\overline{\text{cone}} a) = P(L_a^0) \cdot \left[\prod_{\{z: z \in [\overline{\text{cone}} a] \setminus L_a^0\}} P(z \mid LNz) \right].$$

Abend et al.'s Theorem can be shown to be a special case of this result: The poset X is the set X_{MN} , the level sets have been given previously, the cone closure of any element $a_{i,j}$ is the set X_{ij} , $L_{a_{i,j}}^0 = \{a_{1,1}\}$ for all i,j , and the lower neighbors are as discussed previously.

We next give the result for the simple example defined by Figure 4. This poset has seven elements, $\{a,b,c,d,e,f,g\}$. The diagram depicts the cover relations: e covers c , f covers both c and d , g covers d and b , c covers both a and b , and d covers a . The set of minimal elements is $\{a,b\}$; the set of maximal elements is $\{e,f,g\}$. We see that $X_{\min} = L^0 = \{a,b\}$, and the cover set for L^0 is $L^1 = \{c,d\}$; g is not included in L^1 because it covers d as well, which is not in L^0 . Let us use Theorem 1 to calculate $P(\overline{\text{cone}} g)$. Note that $L_g^0 = \{b\}$; $\overline{\text{cone}} g = \{g, d, b\}$; and $LNg = \{d, b\}$, $LNd = \{a\}$. Thus,

$$\begin{aligned} P(\overline{\text{cone}} g) &= P(L_g^0) \cdot \left[\prod_{\{z: z \in [\overline{\text{cone}} g] \setminus L_g^0\}} P(z \mid LNz) \right] \\ &= P(b) \cdot \left[\prod_{\{z: z \in \{g, d\}\}} P(z \mid LNz) \right] \\ &= P(b)P(g \mid LNg)P(d \mid LNd) \\ &= P(b)P(g \mid d, b)P(d \mid a). \end{aligned}$$

What Theorem 1 shows is that the (marginal) probability of the cone closure of an element can be decomposed multiplicatively into the local conditional probabilities that are assumed in the model.

6. Applications

An application that is immediately apparent from Theorem 1 is that this formula can be used to calculate joint or marginal probabilities used in the generation of textures [Cross and Jain, 1983]. One method often employed generates texture with the same number of pixels at each gray level. Let the current state of values be $\{a_{i,j}\}^{curr}$. A new state of gray values, $\{a_{i,j}\}^{new}$, is generated from the current state by randomly selecting two locations and interchanging their gray values. The scheme to interchange the values uses the Metropolis algorithm [Metropolis *et al.*, 1953] based on the ratio of $P(\{a_{i,j}\}^{new})/P(\{a_{i,j}\}^{curr})$. A decomposition of the form given in Theorem 1 can lead to massive simplification of this ratio. Consider the following algorithm used by [Flinn, 1974]:

```
while not CONVERGED do
  begin
    choose  $a_{i,j}$  and  $a_{h,k}$  with  $a_{i,j} \neq a_{h,k}$ 
     $r := P(\{a_{i,j}\}^{new})/P(\{a_{i,j}\}^{curr})$ 
    if  $r \geq 1$  then
      interchange  $a_{i,j}$  and  $a_{h,k}$ 
    otherwise
      begin
        choose  $\beta \in \mathcal{U}[0,1]$ 
        if  $r > \beta$  then
          interchange  $a_{i,j}$  and  $a_{h,k}$ 
        endif
      begin
    endif
```

7. Conclusions

We have given a more general spatial model that assumes an arbitrary partial order of the underlying set of random variables and given one instance where it could be used. Further research includes generalizing the following properties in [Abend *et al.*, 1965]: 1) multiplicative decomposition for the joint probability of the entire array X when there is more than one maximal element; this results could be used to calculate the joint probabilities as given in the example in section 6. This is Abend *et al.*'s Lemma:

Lemma (Abend et al.). The third-order Markov mesh condition gives a multiplicative decomposition for the joint probability $P(X_{MN}) = P(X)$,

$$P(X) = \prod_{h=1}^N \prod_{k=1}^M P(a_{h,k} \mid a_{h-1,k} a_{h-1,k-1} a_{h,k-1}).$$

The other result we would like to generalize is the formula for calculating the conditional probability $P(a_{i,j} \mid X \setminus \{a_{i,j}\})$ in terms of the local conditional probabilities, where the conditioning set is determined by $\{LN a_{i,j}\}$. This is a generalization of Abend et al.'s second theorem:

Theorem (Abend et al.). The probability of $a_{i,j}$ conditional on the entire array minus $a_{i,j}$ can be shown to be:

$$p(a_{i,j} \mid X \setminus a_{i,j}) = P \left(\begin{array}{ccc} a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} \\ a_{i,j} \mid & a_{i,j-1} & a_{i,j+1} \\ a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} \end{array} \right). \quad (2)$$

Thus, under the given assumptions, this conditional probability requires a conditioning set determined by the eight nearest neighbors. More detailed applications to texture analysis may prove fruitful, as the spatial models provided are more general and not necessarily spatially local.

8. Acknowledgments

The authors wish to acknowledge the support of this research on grants from ONR (N00014-93-1-0001), NSF (DMS-9204521), and NSA (MDA904-92-H-3021).

References

- [Abend *et al.*, 1965] K. Abend, T.J. Harley, and L.N. Kanal. Classification of binary random patterns. *IEEE Trans. on Info. Theory*, IT-11:538-544, Oct. 1965.
- [Cross and Jain, 1983] G.R. Cross and A.K. Jain. Markov random field texture models. *IEEE Trans. on Patt. Analys. and Mach. Intell.*, PAMI-5(1):25-39, Jan. 1983.
- [Flinn, 1974] P.A. Flinn. Monte Carlo calculation of phase separation in a 2-dimensional Ising system. *J. Statist. Phys.*, 10:89-97, 1974.
- [Harary *et al.*, 1965] F. Harary, R.Z. Norman, and D. Cartwright. *Structural Models: An Introduction to the Theory of Directed Graphs*. Wiley and Sons, New York, 1965.
- [Hilbert, 1891] D. Hilbert. Über die stetige abbildung einer linie auf end flächenstück. *Mathematische Annalen*, 38:459-460, Mar. 1891.
- [Metropolis *et al.*, 1953] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equations of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087-1091, 1953.